DOCUMENT RESUME

ED 163 061 TM 008 119

AUTHOR Bierschenk, Bernhard
TITLE A New Approach to Psychometric Problems in the Analysis of Pre-Numeric Data. Didakometry. No. 55.
INSTITUTION School of Education, Malmo (Sweden). Dept. of Educational and Psychological Research.
PUB DATE Mar 77
NOTE 28p.

EDRS PRICE MF-$0.83 HC-$2.06 Plus Postage.
DESCRIPTORS Behavioral Science Research; Cluster Analysis; Computer Programs; *Concept Formation; *Content Analysis; Discriminant Analysis; Experimenter Characteristics; Foreign Countries; *Information Needs; Language Research; Linguistics; Problem Solving; *Researchers; *Research Methodology

ABSTRACT

ANACONDA, a computer-based system for analysis of verbal content by data processing is presented as a scientifically more satisfactory substitute for classical content analysis. This system was applied to 40 behavioral science researchers who were interviewed regarding their channels of information for concept formation, problem discussion, research methodology, and documentation. Researchers' statements were analyzed into a noun-verb noun relationship representing agent, action, and object. The treatment of the researchers' verbal statements by ANACONDA involved an analysis and synthesis of both empirical statements and the relationships between them. These statements were then studied by a discriminant analysis. Finally, an attempt was made to demonstrate the validity of this methodological approach. (Author/CP)

****
* Reproductions supplied by EDRS are the best that can be made *
* from the original document. *
****

ED163061

U.S. DEPARTMENT OF HEALTH, EDUCATION & WELFARE
NATIONAL INSTITUTE OF EDUCATION

THIS DOCUMENT HAS BEEN REPRODUCED EXACTLY AS RECEIVED FROM THE PERSON OR ORGANIZATION ORIGINATING IT. POINTS OF VIEW OR OPINIONS STATED DO NOT NECESSARILY REPRESENT OFFICIAL NATIONAL INSTITUTE OF EDUCATION POSITION OR POLICY

"PERMISSION TO REPRODUCE THIS MATERIAL HAS BEEN GRANTED BY

E. Bierschenk

TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC) AND USERS OF THE ERIC SYSTEM."

A NEW APPROACH TO PSYCHOMETRIC PROBLEMS IN THE ANALYSIS OF PRE-NUMERIC DATA

Bernhard Bierschenk

Bierschenk, B. A new approach to psychometric problems in the analysis of pre-numeric data. Didakometry (Malmö, Sweden: School of Education), No. 55, 1977.

This report suggests a new approach to quantification of the relationships within a whole sentence and presents a system for the analysis of concepts by data processing (ANACONDA). The psychological model for ANACONDA is outlined and a first measure building up concepts with an empirical root is discussed. Panel assessments have been used for quantification resulting in high reliability scores. By means of cluster analysis the empirical material concerning researchers' views of scientific information and documentation has been homogenized. On the basis of the results obtained, the relation patterns have been studied by means of a discriminant analysis. Finally an attempt has been made to demonstrate the validity in the methodological approach used.

Keywords: Psycholinguistics, concept formation, cognition, interview data, psychometrics, information and documentation, regression analysis.

CONTENTS

	Sid
1. BEHAVIOURAL SCIENTISTS' PERCEPTION AND DEFINITION OF RESEARCH PROBLEMS: SOME RESULTS AND CONCLUSIONS	2
2. FORMULATION OF RESEARCH PROBLEMS: DEVELOPMENT OF A MODEL FOR INVESTIGATION	4
2.1 Development of a model for investigation	4
2.2 Some methodological prerequisites	5
2.2.1 Researcher population and random sample	5
2.2.2 Data collection and data analysis	5
3. DEVELOPMENT OF A STRATEGY FOR THE ANALYSIS OF VERBAL DATA	7
3.1 Development of a computer-based content analysis	7
3.1.1 Control of coder agreement	8
4. A PSYCHOLINGUISTIC PROCESS MODEL	10
4.1 Psychological paradigm and unit of analysis	10
4.1.1 Manifest level: Basic elements and syntax	10
4.1.2 Latent level: Cognitive structures	10
4.1.3 Inference: Reconstruction of reality	11
4.2 Operationalization of the psycholinguistic process model	11
4.2.1 Scaling of adjectives and verbs	12
4.2.2 Generation of concepts	13
5. RECONSTRUCTION OF REALITY: MAIN RESULTS FROM AN ANALYSIS OF DISCRIMINATION	15
6. VALIDITY OF ESTABLISHED STRUCTURES	20
7. REFERENCES	22

1. BEHAVIOURAL SCIENTISTS' PERCEPTION AND DEFINITION OF RESEARCH PROBLEMS: SOME RESULTS AND CONCLUSIONS

This report presents our method for the analysis of concepts by data processing (ANACONDA). The method has been developed in connection with an empirical investigation of the way in which researchers perceive, structure and define problems suitable for behavioural science research. The research has been financed by the Swedish Board of Education.

ANACONDA is an attempt to develop a method for the analysis of content that is scientifically more satisfactory than classical content analysis methods. ANACONDA has been applied to statements made by researchers on scientific documentation and information. A few examples of results from this study will be given. Following Miller, Galanter & Pribram's (1960) "Plans and the structure of behavior", we can say that the researcher's plan for solving his information problem contains intentions and goal notions, plus an idea of which means can be used to achieve goals, i.e. means-goal hierarchies. The intention is to get in principle two types of information (1) for demarcation of concepts and (2) about research methods.

When the intention is to get information for the demarcation of concepts the strategy (means) used is problem discussion (discussion seminars, project meetings, informal literature seminars). On the other hand, when the intention is to find information concerning research methods the strategy developed is to a certain extent bibliographic information-seeking and visits to international symposia. But primarily tutors and fellow-researchers are asked. The conclusion we can draw is that researchers do not use the same strategies for getting information about demarcation of concepts and information on research methods respectively. Furthermore since both types of information are negatively related to each other, we can conclude that the information-seeking strategy used is a function of the type of information sought.

Instrumentalization, i.e. the technical systems available for channelling information, is used to a certain extent to obtain bibliographic references, i.e. information about information. But the actions that form the building stones of the researchers' information-seeking strategies express a neutral to negative attitude.

One condition ANACONDA must fulfil is that the method must lead to a valid reconstruction of the cognitive structures which we assume underlie verbal expressions. An attempt to demonstrate the validity of the method is made in the last Chapter.

Considering the scientific debate of recent years on process research and the

marked limitations of various assessment schedules as data-gathering methods, a continuous development of ANACONDA should represent a significant contribution to the research methods that are available to behavioural scientists.

The purpose of this introductory presentation of research results achieved by applying ANACONDA to behavioural scientists' verbal statements concerning scientific information and documentation is (1) to motivate the reader to expend some energy in following through this exposition and (2) to offer some pointers to the theoretical framework behind the development of a method for studying complex psychological phenomena, such as communication by means of symbols.

2. FORMULATION OF RESEARCH PROBLEMS: DEVELOPMENT OF A MODEL FOR INVESTIGATION

Psychological research has investigated how we optimally may choose between behaviours when a problem is "well defined". On the other hand, we know very little about how the process of obtaining "well defined problems" develops in spite of the fact that knowledge of the development of problem formulation processes is especially important in research aiming to find solutions to original problems which can be attacked by means of scientific methods. The problem defined must also contain clues pointing to the possibility of gaining new knowledge. The researchers' ability to observe and to gain insight should be regarded as essential factors in the process of problem perception and problem definition. It is namely in no way self-evident that ideas that are fruitful from a scientific point of view are included in the "official" lists of experienced problems that are sometimes compiled by consumers of behavioural research. Fruitful and convincing ideas are usually a result of continuous and tenacious intellectual work covering a period of years.

2.1 Development of a model for investigation

Even though the purpose of an empirical investigation is to find patterns in a set of empirical data, the first step in carrying out such an investigation should be to discern important components which should be included in a model intended to reflect adequately the problem formulation process.

In 1972 a pilot study was initiated to find indications of (1) how problem perception and definition together with the evaluation of problems develops as felt by behavioural scientists and (2) how the information search process influences the problem formulation process and vice versa. In this study we focussed our attention on the initial phase of the research process. This developmental process starts with the preparation of an outline of a research project and ends with the formal decision to give the outline project status and start developing a more detailed project design. Guided by the results of this study we developed a model on which we could base an interview study. In 1973, that investigation was carried out in order to collect a representative set of data about Swedish researchers' strategies of behaviour during the initial phase of a research process. The model which guided the investigation contains the following components:

1. The researcher's motivation, i.e. his own aims and intentions.
2. The researcher's perception of problems, i.e. generation of ideas, norms and values.
3. The researcher's choice, transformation and structuring of problems and generation of hypotheses.

4. The researcher's choice of strategies concerning the search for information, the design and implementation of research plans.
5. The researcher's choice of methods and techniques for information and documentation of research results, design of investigations, collection and processing of data.
6. The researcher's frame of reference, i.e. the extent to which the researcher identifies with a specialization, with a particular discipline, with an interdisciplinary grouping or invisible colleges.
7. Research organizations, both formal and informal within which researchers work.

A detailed description and an attempt at quantification and systematic analysis of the model have been presented in B. Bierschenk (1974). An English summary of the main results is given in B. Bierschenk (1977 a).

2.2 Some methodological prerequisites

The use of a given research strategy and certain special techniques for data collection, processing and analysis always implies a (hopefully) conscious choice. But it also means a focussing on or systematic search for new knowledge and a decision on methods, which aims at reducing as far as possible the influence of various conceivable sources of error.

2.2.1 Researcher population and random sample

The research results that will be presented below are based on forty randomly selected researchers out of a population of 126. The following criteria must be fulfilled for a person to be considered a researcher belonging to the population of investigation. A researcher is a person who (1) has studied at least one of the behavioural sciences (education, psychology, sociology) or statistics, (2) has taken at least a licentiate degree, (3) is connected with an on-going, recently started or completed project, seen from the perspective of 1972, (4) belongs to or is connected with at least one of the (then) eleven departments of educational and psychological research in Sweden, (5) is resident in Sweden and (6) has been included in one of the catalogues listing social science research of staff employed at the departments. These catalogues must have been published during the period 1967-1972. Catalogues presenting developmental work are not counted.

For a detailed description of the population and random sample and for an account of the data collection see B. Bierschenk (1974).

2.2.2 Data collection and data analysis

If we assume that it is only individuals that can formulate problems of scientific interest, it becomes necessary to let individual persons express themselves freely. We chose to use an interview method for the collection of data. Our assump-

Box 1. Data processing and analysis

Processing Models/techniques	Analysis Models/techniques	Description Inference
Impressionistic content analysis Annerblom (1974)	Dichotomization of attributes irrespective of frequency	Descriptions with the purpose of formulating hypotheses and discovering new unique attributes
Computer-based content analysis Bierschenk & Bierschenk (1976)	Strict or rigid counting of frequencies with the purpose of applying statistical models	Objective, reliable and detailed description with the purpose of testing hypotheses. Applying multivariate models for data analysis
Formalized statements together with assessment scales B. Bierschenk (1974 and 1977 a)	Index construction. Analysis of frequency distributions by means of statistical models	Descriptive and inference statistics. Description by means of m, s, r, F-ratio, ω^2

tion was that the researcher's chance to give free and unresticted answers would provide information of high validity, at least as long as we could assume that the researchers were seriously interested in participating in the study. But we also constructed an assessment schedule for the purpose of (1) getting the researchers' views on certain aspects in a formalized form, (2) getting the researchers' possible suggestions for improvements and comments on wordings and statements in the schedule. The scheme for data processing and analysis is presented in Box 1.

When the researcher demonstrate which aspects are to be mapped, he often constructs models and data matrices, in which the rows usually represent the measuring objects (in our case researchers). When using psychological tests or assessment scales, one gets scores that can be used directly for setting up data matrices. Such scores are not obtained immediately, however, when the basic material consists of verbal data. Thus it will be necessary to transform this type to numerical values. But the analytical problem in making inferences from verbal expressions is much the same as with all other types of raw data. We have to infer specific events, behaviours or characteristics which are associated with the objects of an investigation.

3. DEVELOPMENT OF A STRATEGY FOR THE ANALYSIS OF VERBAL DATA

If we want to overcome successfully the difficulties that complex problems pose for researchers in the behavioural sciences, a great deal of openness is required and readiness to use new research strategies. But these should be precisely defined and explicitely described.

The development of ideas, the formulation and solution of problems are behaviours that are closely connected with man's specific ability to express himself verbally. Attempts at exploring and studying not only problem-solving behaviours but also the complex psychological process behind such behaviours are not confined to the field of psychology. For the past twenty years or so such experiments have also been carried out within several other branches of science, such as mathematics, artificial intelligence, information processing and quantitative linguistics. The common factor of the research work done within the various disciplines is the goal of investigating and giving form to invisible symbolic processes and mechanisms. Instead of using simple mathematical formulas to try to describe such cognitive operations, researchers working in the fields named above are trying to develop computer programmes to describe them and to test models and theories on what complex psychological structures look like and how processes develop.

These attempts at developing a theory on the content of the messages by means of which people communicate with each other aim at increasing our understanding of the cognitive structures that are assumed to form the basis for a human being's verbal expressions. By developing our assumptions step by step and continuously testing them, we try to determine their validity. The main question we have asked ourselves is the following:

Can we by means of numerical analysis and quantitative description identify and categorize cognitive structures in pre-numeric data, such as interview texts?

3.1 Development of a computer-based content analysis

It is typical for written or spoken text that it is of great complexity and that the type of information to be extracted from the material is seldom or never collected in one single place in the text. If structural relationships are to emerge all the same, the text must be prepared on the basis of certain assumptions. In Bierschenk & Bierschenk (1976) a flow chart is presented stating the individual steps in the development of the method for computer-based content analysis that we suggest.

The chances of ready-developed methods being applicable often depend on the

appearance of the material. The attempts at analysis that have been described in the literature and that are of interest to our analysis have been developed with written text as a basis. But since our material is spoken language text (transcribed from recording tape), which when uttered was meant for the ears of the interviewer alone, it became necessary to build up our own system of rules and codes. A preliminary manual and some test results are presented in I. Bierschenk (1974), but a new version will be published during 1977.

The theories of verbal representation which we have found most interesting for our analysis are Schank's (1972) "Conceptual Dependency Theory" in connection with "Case Grammar". In Bierschenk & Bierschenk (1976) we have shown how a model containing symbolic representations of concepts and conceptual relation can be applied to empirical materials. In our presentation two kinds of dependence and independence are distinguished: vertical relations, i.e. between several independent concepts. We also discuss relations, i.e. the relation of clauses to each other.

We take the theory of necessary parts into account in another way too. Since language is economical in relation to the thoughts behind the utterance, not all necessary parts are included in a speaker-listener situation. The necessary concept apparatus exists in the listener and a syntactically incomplete sentence is understood all the same. But the computer does not have this understanding, so therefore we supplement in the parts that are missing. It is our task to code complete complexes and complete Subject-verb-Object paradigms. Coding rules have been worked out and rules for the treatment of text for input in the computer have been specified. The way in which the feeding takes place, how identification is specified in the coding and the way in which we build up our dictionary base are described in Bierschenk & Bierschenk (1976).

3.1.1 Control of coder agreement

By using the concept "Computer-based content analysis" we wish to point out that we do not intend to develop a method for automatic text analysis. At the same time this means that the basic material must first be coded before mechanical processings of various kinds can be carried out. The success with which two independent coders have been able to apply different coding rules in a similar way is described in detail in Berg (1974). The intercoder agreement was examined with regard to (1) segmentation of concepts, (2) segmentation of clauses, (3) assignment of codes to concepts and (4) assignment of codes to themes. All the comparisons are of the same type, i.e. either there is agreement or there is not.

//

The index values for these four steps vary for comparable text between i =. 82 and i =. 98 (i: Osgood's index for agreement). More detailed information and test results obtained by applying the binomial test to the data may be found in Bierschenk & Bierschenk (1976).

4. A PSYCHOLINGUISTIC PROCESS MODEL

It is very difficult to try to map what is really meant in the research literature by content analysis, since each content analysis technique is based on a specific way of regarding the content in a message. A content analysis presupposes that we can define what is to be measured and counted in the analysis.

In our discussion we have now reached the point where we have to introduce a psychological paradigm. We need a paradigm that takes into account not only representation of relations between concepts, but also goal-directed behaviour. The paradigm in our psycholinguistic model which serve this function is the well--known Agent-action-Object (goal) paradigm. The central importance of the directed actitivy can be seen, among other things, by the function of the action in determining the nature of the AaO paradigm.

4.1 Psychological paradigm and unit of analysis

The components of the AaO paradigm are briefly defined as follows:

Agent. Centres of action or goal-seeking entities that make use of resources to reach their goals. This definition also includes groups, organizations or abstractions that fill the function of being an agent.

Action. A direct action that is carried out by an agent for the purpose of achieving a goal. The action defines the content of the AaO paradigm.

Object. Everything that an action can be directed at or implemented with.

By means of the AaO paradigm, the components that form a natural content, i.e. an observable sentence, are isolated. By this is meant the fundamental form for a statement that is expressed by the $noun_1$-verb-$noun_2$ relationship. The analysis and synthesis of the psychological phenomena to be studied is constructed from this "whole observation sentence" (see Quine, 1972).

4.1.1 Manifest level: Basic elements and syntax

While agent and object (noun) are specified by attributes (adjectival phrases), the verb states the relation between the nouns, i.e. actions, events or state. The order of these basic elements is stated by means of syntax. By using empirically constructed "dictionaries" and a system of rules we hope to be able to construct theories and models that can be used to describe strategies of actions and events.

4.1.2 Latent level: Cognitive structures

Psychological research, and in particular its psycholinguistic branch, has long been trying to map the psychological processes underlying linguistic sentences. In explaining the acquisition of language the process theoreticians assume a cog-

nitive mechanism. Thus it is assumed that there are regularities in complex phenomena that can be observed or predicted. In our model we symbolize the individual prerequisites of a researcher by means of theoretical constructs that are arranged around the Test-Operation-Test-Exit (TOTE) paradigm (see Miller, Galanter & Pribram, 1960). TOTE symbolizes the many cyclical processes which are assumed to govern the organism's selection of information. In the same fashion we assume that TOTE at the cognitive level governs the choices and decisions necessary for the development of a research process.

Problem formulation and problem solving are fundamental human operations. But since science cannot be regarded as a determinable object or a determinable set of problems, but must be considered as a way of attacking problems (methods and goals), everything can be changed into an object for scientific examination.

Thus in the model method is defined as the process of a systematic search for new knowledge. This process is made explicit by research designs or designs for a systematic information search. In the model, goal is defined as the focusing of attention on the objectives of a piece of research. Plans are made concrete by the research strategies established. By strategy is meant a sequential or hierarchical order or a set of instructions which steers the actions of an organism according to a plan, i.e. the performance. Transferred to the problem of searching for scientific information, that means the use of such advanced systems as computer-based I&D systems. Thus in the model it is the method that defines the content in the Problem-Method-Goal (PMG) paradigm.

On the latent level it is (1) generally available knowledge (evidence) and (2) anticipated problem solutions (plans and strategies) that form indispensable elements in the research process.

4.1.3 Inference: Reconstruction of reality

In our psychological model, the operationalizing and interpreting functions are summarized under the term "concept". It is assumed that every utterance is based on concepts that form the basis for key words in a clause. We hope that in describing the research process the concepts will reflect (1) the researcher's motives and (2) frame of reference, which are assumed to steer his (3) formulation of hypotheses and (4) development of designs, (5) choice of methods and (6) techniques or instruments necessary for the accomplishment of a scientific task.

4.2 Operationalization of the psycholinguistic process model

The building up of concepts in a given context presupposes a system of rules that state how and in what order different elements are to be linked together. We assume, for example, that the relations that exist between nouns and adjectives and

between nouns and verb represent the same relations as those between empirical phenomena. If, for example, we want to state that a set of nouns is modified by a set of adjectives we can formalize this relationship.

The usual way of stating a relation is to specify a rule that says what are to be regarded as elements, pairs etc. If, for example, we wish to state for our sample space ($\mathcal{S}$) that A is the set of all nouns and B is the set of all adjectives modifying A, this relation can be stated more formally in the following way:

$$\mathcal{S} = \left\{ (a, b) \in A \times B \parallel a \text{ modified by } b \right\}.$$

Further, we assume that nouns get part of their empirically specified content through adjectives and verbs to which they are linked. By scaling adjectives and verbs we can acquire numeric descriptions and quantitative analyses of text. When we have in this way observed similarities or co-variations and defined different properties, we can carry out multivariate analyses in order to determine the position of a particular property in a number of latent dimensions.

It should also be mentioned that work is under way to find out how we might quantify the "stretching property" of adverbs (see Cliff, 1969, p 144). The measurement suggested by I. Bierschenk (1977) implies that we eventually may be able to describe the whole vertical relationship, i.e. between dependent and independent concepts.

4.2.1 Scaling of adjectives and verbs

Researchers have usually used semantic differentials linked to adjectives when describing an object. We consider that the boundaries between adjectives and verbs are vague and that in principle all "dependent" concepts should be utilized in the description of a phenomenon (see Ross, 1969). Since in our analysis we take into consideration "syntactic behaviour" and regard both adjectives and verbs as descriptive concepts, we have decided to scale them. An adjective describes a noun directly, while the verb has the same function more indirectly. The scaling was carried out by means of seven-point assessment scales, the bipolar end-points of which are described as pairs of adjectives (1) positive-negative, (2) active-passive and (3) strong-weak. A detailed account of this is to be found in Bierschenk & Bierschenk (1976).

Panel assessment is the method of approach which was used to produce assessment values for all the adjectives (n = 570) and verbs (n = 883) extracted in the dimension concerned (8 718 assessments). The random sample of 15 persons included in the assessment panel is covered by our definition of "researcher", which should make it possible to generalize the results.

In order to avoid, among other errors of measurement, some concepts being

liable to a fatigue effect, the order of the words has for each individual assessor been determined at random, i.e. 15 different random orders of sequence were generated by the computer. In addition the three dimensions have been separated in order to counteract any mixing of the individual scales. This means that each person received six different random orders of sequence. The assessment panel made its assessments during the spring term of 1975.

Since we shall be making use of the mean assessment of the assessors and all adjectives and verbs respectively that define a particular noun, we have not only estimated intraclass correlations but also the structure of the assessment (see Guilford, 1954, pp 253-254). In order to obtain a coefficient for maximal reliability for the respective scales, each position on the seven-point scale has been weighted according to the component analysis. The coefficient for maximal reliability was introduced by Lord (1958). This coefficient is a simple function of the largest characteristic root of correlation matrix for the variables forming the scale. A detailed account of our results in the reliability studies carried out is presented in Bierschenk & Bierschenk (1976). In Table 1, therefore, only the most condensed reliability estimations will be given.

Table 1. Reliability in the assessments of adjectives and verbs (α_{max})

Object of assessments	Dimension (1) Evaluation	(2) Activity	(3) Power
Adjective	.97	.92	.88
Verb	.95	.93	.86

In order to create a weighted summation variable, the weights are used from the first unrotated component. Each assessment is multiplied by the weight for the respective assessor and dimension. Then the sums of these products were calculated. In order to create weighted means for each adjective and verb, the totals have been divided by the sum of the weights. If any assessment has been dropped, the corresponding weight has been subtracted when the sum of the weights was calculated.

4.2.2 Generation of concepts

In the behavioural sciences text analyses are usually based either on an examination of the text, looking for specific key concepts, or on the frequency of a specific concept in the text. The application of a concept system in the coding of texts assumes that the latent structure of the text is reflected in the concepts and in the structure represented by a system of categories. A content analysis based on

dichotomous decisions about or frequency distribution of concepts can, however, prove to be insensitive to the researcher's own terminology and way of structuring text.

Treatment of researchers' verbal statements by ANACONDA involves an analysis and synthesis of both empirical statements and the relationships between them. This means that we cannot be content with a traditional lexico-graphic listing of words as a base for an approximation of the researcher's implicit models for the research process, but have to build up dictionaries empirically. At the same time it also means that we must be able to handle large amounts of data so that meaningful statistical descriptions and analyses become possible. The way in which we have carried through the task is mapped by the following 15 step analysis programme:

Step 1: An analysis of how many agents the 40 researchers have produced in four questions out of 51.

Step 2: An analysis of how the agents are distributed over the researchers.

Step 3: An analysis of how many objects the 40 researchers have produced in four questions out of 51.

Step 4: An analysis of how the objects are distributed over the researchers.

Step 5: A cluster analysis of the agents: (1) BMD P01M, in which the agents are treated as variables and (2) BMD P02M, in which the agents are treated as measuring objects (Dixon, 1975).

Step 6: A cluster analysis of the objects: (1) BMD P01M, in which the objects are treated as variables and (2) BMD P02M, in which the objects are treated as measuring objects.

Step 7: A cluster analysis of blocks: BMD P03M, in which agents and objects respectively form one block and the researchers another and where the threshold values (step lengths) are placed at .20 and .10 respectively.

Step 8: An analysis of the agents' coincidences with objects, where the coincidences are determined through verbs

Step 9: An analysis of the number of agents that occur at least five times in the interview material.

Step 10: An analysis of the number of objects that occur at least five times in the interview material.

Step 11: An analysis of the distribution of the number of agents that coincide with such objects as occur at least five times.

Step 12: An analysis of distribution of the number of objects that coincide with such agents as occur at least five times.

Step 13: A cluster analysis of agents and objects respectively: BMD P01M, in which (1) agents and (2) objects are treated as variables.

Step 14: An analysis of the agent clusters' coincidence with the object clusters, where the coincidences are determined by verbs.

Step 15: A discriminant analysis of reduced coincidence matrices: The computer programmes used are MANOVA & DISCRIM and the programme from SPSS (Nie et al., 1975).

5. RECONSTRUCTION OF REALITY: MAIN RESULTS FROM AN ANALYSIS OF DISCRIMINATION

The relations that are assumed to exist between concepts are implicative or inferential and they are intended to be operationalized by means of analysis models based on perceived co-variations or correlations. In order to be able to study the dimensionality in the researchers' verbal statements, we have applied the discriminant analysis model. A summary of the main results will now be given. More detailed information and a discussion of how the prerequisites of the model have been fulfilled may be found in B. Bierschenk (1977 b).

The assessment scores describing the interview material have been arranged in accordance with the following scheme of co-variation:

G: Scales 1(1)3; A: Agents 1(1)14; C: Concepts 1(1)6.

Working with the complete AaO relation, we have defined the derived "clusters of agent" as measuring objects and the "clusters of concept" as variables. These together with the scales form three groups for assessment. The purpose of this co-variation scheme is to study the linear combinations that must be formed if the variance between the scales is to be maximized. A discriminant analysis of this type means that we study the co-variance between the scales. In this analysis we study the importance of objects and scales for a particular agent.

When three groups exist, two discriminant functions can be formed. The discriminating power and test statistics are given in Table 2 together with the standardized functions.

Table 2. Discriminating power of discriminant functions and standardized discriminant function coefficients for concept clusters

Function	Discrimination power λ	%	R	Statistical tests Λ	χ^2	df	p
1	1.36	75.85	.76	30	44.98	10	.00
2	.43	24.15	.55	70	13.28	4	.01

Concept clusters	f_1	f_2
1. Bibliographic reference	.50	.04
3. Discussion of problems	.44	.82
4. Channnels of information	.53	-.71
5. Information on research methods	-.17	-.49
6. Information on demarcation of concepts	.32	.33

λ Eigen value
% Relative percentage of total variance
R Canonical correlation
Λ Wilk's lambda
χ Chi square
df Degrees of freedom
p Level of significance

As can be seen from Table 2, the first function is much more important for the separation than the other. But both are successful in separating our three scales. The standardized coefficients reproduced in Table 2 can be interpreted in the same way as factor loadings or beta weights in a multiple regression analysis. In this sense the coefficients state which clusters contribute most to differentiation in the respective dimension.

The clusters that are important for the first function are "Bibliografic references" and "Channels of information" which in addition show high negative weight in the second function.

The ones that are important for the second function are "Discussion of problems" and with a negative sign "Information on research methods". "Information on demarcation of concepts", on the other hand, is of little importance for both functions. This result is in perfect agreement with the results presented in B. Bierschenk (1974). The omitted second concept cluster "Research organization" has no discriminating power, but rather a reducing effect when the cluster is combined with the others.

A summary of how far we have succeeded in separating the scales and the extent to which the classification of the individual agent clusters has been achieved is shown graphically by the two-dimensional discrimination space in Figure 1. The centroids (*) in Figure 1 state the mean value of the discriminant scores for each scale and respective function. As can be seen from Figure 1, Function 1 discriminated well between "Evaluation" (1) and "Power" (3). Function 2 is needed to distinguish as far as possible "Activity" (2) from "Power" (3). Figure 1 shows that it is easier to differentiate between evaluation and power than between evaluation and activity, or power and activity. But the separation of function 2 would probably emerge even more clearly if it were not for three outliers.

The agent cluster "Dissemination of information" with the discriminant scores (1.63, -1.61) in the evaluation dimension and (-1.67, -2.60) in the activity dimension is the one that deviates markedly from the other clusters. The other agent cluster that falls outside its group affiliation is "Methodological problems" with the discriminant scores (1.15, 2.00) in the activity dimension

These circumstances can be studied in more detail by examining the linkages of agent clusters with regard to the values presented in Table 3.

The agent cluster "Dissemination of information" that includes the agents "Literature" and "Symposium" (see B. Bierschenk, 1977 b, Box 1) will be used to give an example of such an examination. If "Dissemination of information" is related to "Bibliographic reference" (that encompasses the concepts "Literature", "Reference" and "Journal") via the verb linkages, it proves that the <u>actions</u> ex-

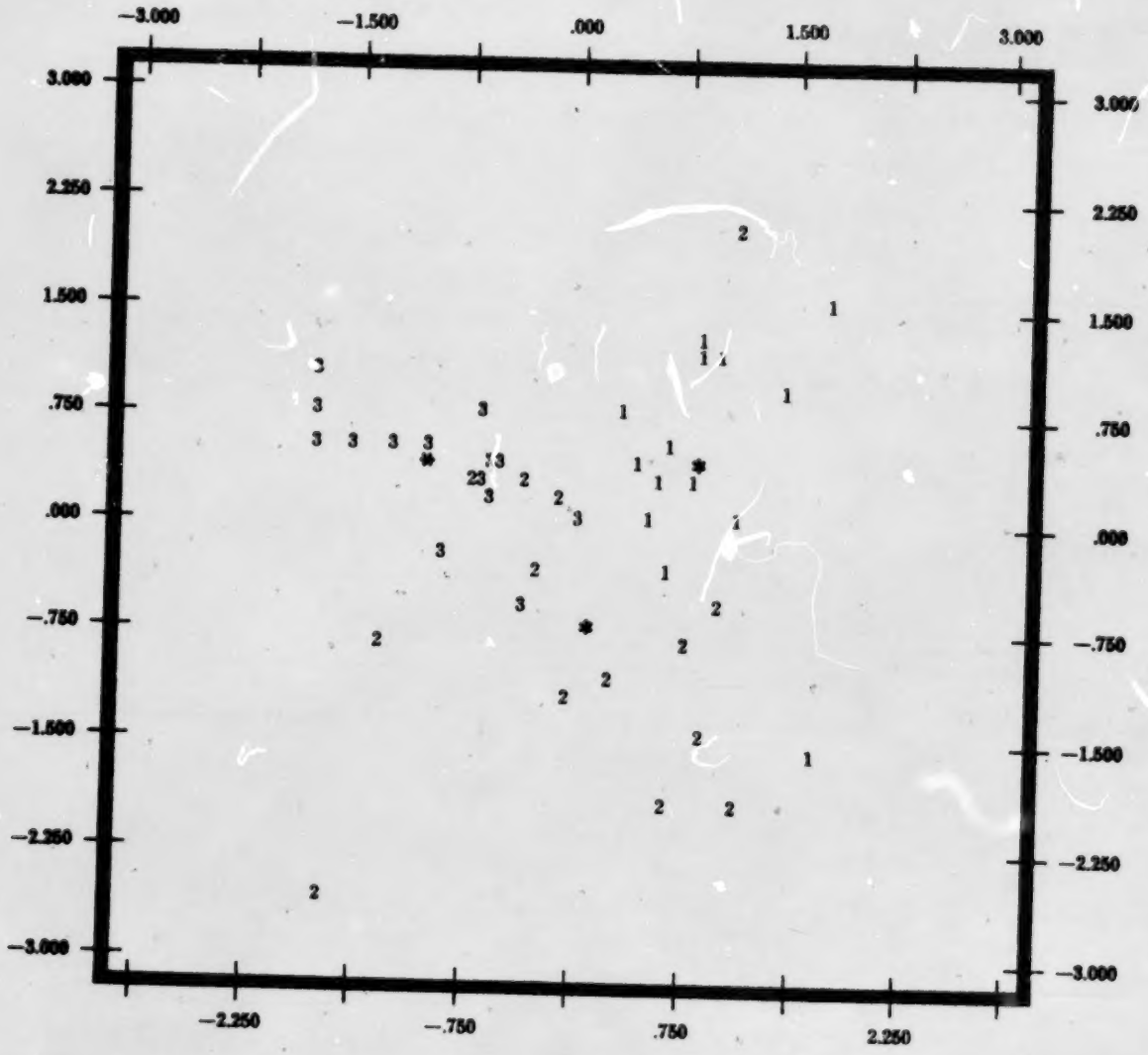

Figure 1. Plot of discriminant score 1 (horizontal) vs. discriminant score 2 (vertical). * Indicates a group centroid.

Table 3. Mean values and standard deviations for 14 agent clusters, 6 concept clusters and 3 assessment dimensions

	Bibliographical reference			Research organization			Discussion of problems			Channels of information			Information on research methods			Information for demarcation of concepts		
	1	2	3	1	2	3	1	2	3	1	2	3	1	2	3	1	2	3
1	4.12	3.65	3.77	4.65	5.34	4.64	4.31	3.69	3.78	4.34	4.78	4.21	3.09	2.95	3.76	4.75	3.23	3.77
2	4.20	4.54	3.87	4.17	4.00	4.05	4.60	4.00	3.86	4.23	4.55	3.92	4.50	5.15	4.15	4.38	4.37	4.19
3	4.29	4.21	3.78	4.17	4.00	4.05	4.55	3.50	3.76	4.18	4.06	4.00	3.35	2.37	3.73	4.03	3.17	3.52
4	4.18	4.60	4.01	4.17	4.00	4.05	4.77	3.73	3.95	4.11	3.89	3.84	4.06	4.11	3.86	4.28	4.19	4.04
5	4.18	4.63	4.23	4.17	4.00	4.05	4.06	2.59	3.37	4.05	4.09	4.02	3.56	3.66	3.80	5.35	5.05	5.04
6	4.38	4.19	3.84	4.43	4.53	4.01	4.23	4.84	3.85	3.97	4.01	3.95	4.10	4.16	3.83	4.39	5.28	4.22
7	4.24	4.37	4.01	3.85	3.23	3.78	4.31	3.69	3.78	4.07	4.27	3.92	4.43	4.63	3.98	4.52	5.23	4.11
8	4.65	5.63	4.27	4.17	4.00	4.05	4.31	3.69	3.78	4.15	3.88	3.55	4.95	5.71	4.60	4.20	4.24	3.77
9	4.35	4.46	3.91	4.37	4.45	4.06	4.30	4.27	4.09	4.11	4.39	3.94	4.35	4.27	3.97	4.32	3.98	3.77
10	4.15	4.15	3.86	3.91	3.59	4.36	4.31	3.69	3.78	4.23	4.51	3.82	4.06	4.11	3.88	3.93	4.57	3.99
11	4.59	3.21	3.32	4.03	3.21	4.07	3.46	2.39	3.75	4.60	4.21	3.76	4.06	4.11	3.88	5.22	4.57	3.82
12	4.50	2.46	2.77	4.17	4.00	4.05	4.86	4.71	4.14	4.17	4.20	3.86	4.23	4.84	3.85	5.22	4.57	3.82
13	4.81	3.93	3.99	4.59	5.22	4.36	4.06	2.59	3.73	4.22	3.93	3.73	4.07	3.25	3.42	4.57	4.23	3.95
14	4.78	4.66	4.07	4.23	2.39	3.14	4.23	4.24	3.53	4.02	4.04	3.69	4.03	4.19	3.69	4.06	2.59	3.37
m	4.38	4.19	3.84	4.17	4.00	4.05	4.31	3.69	3.78	4.18	4.20	3.86	4.06	4.11	3.89	4.51	4.23	3.96
s	.25	.77	.40	.39	1.05	.45	.41	.89	.26	.16	.28	.19	.54	.99	.30	.48	.82	.41

Agent Clusters

1 Social-psychologically oriented researchers
2 Pupil-oriented researchers
3 Language-oriented researchers
4 Science-oriented researchers
5 Cognition-psychology-oriented researchers
6 Researchers interested in methodological problems
7 Researchers interested in programmes for applying influence
8 Linguistically-oriented researchers
9 Humanistically-oriented researchers
10 Researchers interested in sosialization
11 Dissemination of information
12 Type of information and transference of information
13 We
14 Person

Dimensions

1 positive/negative
2 active/passive
3 strong/weak

press a weak positive evaluation (m = 4.59), passivity (m = 3.21) and weak power (m = 3.32).

The second relation to be studied is between "Channel of information" (including e.g. "Library" and "Psychological Abstracts") and "Bibliographic reference", which expresses much the same evaluation (m = 4.60) but higher activity (m = 4.21) though with weak power (m = 3.76).

The third relation of interest is between "Channel of information" and "Problem discussion" (containing the concepts "Discussion" and "Problem"). The actions expressed by the verb show a negative evaluation (m = 3.46), passivity (m = 2.39 and weak power (m = 3.75).

The fourth relation to be investigated is between "Dissemination of information" and "Information on Research Methods" (including e.g. "Design" and "Measurement"). The relating verbs express actions that are somewhat more positive (m = 4.06) in their evaluation. They are also somewhat more active (m = 4.11) and show a marginal increase in power (m = 3.88).

Against the background of these results, the following conclusions can be drawn. The dissemination of information by means of literature and symposia do not contribute noticeably to problem discussion. Moreover this agrees entirely with our expectation that the activities connected with symposia do not express much dynamic, i.e. activity and power. Since this is the case, the actions should not express any very great positive evaluation either. The actions that associate dissemination of information via literature and symposia with information on research methods are rather neutral in all three dimensions. But it is also plain that information on research methods is to a somewhat greater extent sought and disseminated via literature and symposia than information that is of importance for problem discussions. This result coincides perfectly with the researchers' statement that they want a change towards a greater depth in research methods as reported in B. Bierschenk (1977 a).

6. VALIDITY OF ESTABLISHED STRUCTURES

Instead of the "Discussion" or "Summary" usually closing a presentation, we wish to conclude our account of methodological developments by trying to demonstrate the validity of the results we have arrived at with ANACONDA.

The researchers' answers to questions concerning information and documentation have been systematized by Annerblom (1974) in the form of an impressionistic content analysis. This analysis and the evaluation on seven-point bipolar scales presented in Bierschenk (1974, 1977 a) will be used to demonstrate the validity of the ANACONDA-method.

The results show that a thorough and systematic check of research publications is the exception rather than the rule. Personal contacts are felt to be the best source of information, but do not appear to play an important part in the researcher's attempts to bring about a mutual exchange of information. The library is in many cases used because of good personal relations with the library staff. But the expectations of obtaining information from libraries are low, however. It emerges from the comments on the evaluation of reference organs that the researchers feel dissatisfied with their quality and wish for improvements. Symposia are attended roughly once a year. The information obtained from reference organs and symposia is evaluated lowest. The methods used in searching for references to literature are unsystematic and are employed periodically.

The researchers do not appear to search for information on research methods while the problem discussion is underway or when information for the demarcation of concepts is sought. This result is also supported by all the critical opinions on printed information material. Nor is information on research methods that is available in handbooks and works of reference sought to any great extent, instead such information is sought mainly from tutors and fellow-researchers.

The impression given by this account and the conclusion reaches in connection with the outcome of the discriminant analysis are utilized in Box 2 to make explicit how researchers perceive information problems and which strategies they implement to reach the scientific information needed for a successful definition of research problems.

The result pictured in Box 2 represent a stereotyped account and therefore need some illumination: The researcher seeks information for demarcation of concepts mainly via problem discussions. Information on research methods, on the other hand, is sought neither by the use of different channels of information nor through problem discussion. Information for demarcation of concepts seems to be a particular type since it is negatively related to information on research methods. But the evaluations, the activity and the power that the actions express show neutral to negative attitudes. 34

Box 2. Search for scientific information

Strategy	(1) demarcation of problems	(2) information on research problem
Problem discussion	+	-
Bibliographic reference systems	+	0
Consultation of tutors, fellow researchers and symposiums	-	+

+ strategy is used
- strategy is not used
0 strategy is used occasionally

7. REFERENCES

Annerblom, M-L. En impressionistisk innehållsanalys av intervjuer med forskare på pedagogiska institutioner i Sverige. /Interviews with researchers in deparments of education in Sweden: An impressionistic analysis. /Pedagogisk-psykologiska problem, No. 255, 1974. /In Swedish/

Berg, M. Reliabilitetsprövning av en metod för innehållsanalys av intervjutext. /Reliability testing of a method of content analysis to interview texts. /Testkonstruktion och testdata, No. 26, 1974. /In Swedish/

Bierschenk, B. Perception, strukturering och precisering av pedagogiska och psykologiska forskningsproblem på pedagogiska institutioner i Sverige. /The perception, structuring and definition of educational and psychological research problems at the departments of education in Sweden. /Pedagogisk-psykologiska problem, No. 254, 1974. /In Swedish/

Bierschenk, B. Research planning from a micro-ecological perspective: Summary of interview study. Educational and Psychological Interactions, No. 60, 1977. (a)

Bierschenk, B. A computer-based content analysis of interview texts. Numeric description and multivariate analysis. Didakometry, No. 53, 1977. (b)

Bierschenk, B. & Bierschenk, I. A system for a computer-based content analysis of interview data. (Studia Psychologica et Paedagogica, 32) Lund: Gleerup, 1976.

Bierschenk, I. Konstruktion av ett regelsystem för en datorbaserad innehållsanalys av intervjutext: Preliminärmanual och några utprövningsresultat. /Construction of rules for a computer-based content analysis of interview texts: A preliminary manual and some evaluation data. /Testkonstruktion och testdata, No. 25, 1974. /In Swedish/

Bierschenk, I. En studie av modifierande ords kontextuella funktion: En principdiskussion och några kvantifieringsförsök. /A study of the contextual function of modifiers: Foundations for quantification. /Pedagogisk-psykologiska problem, No. 314, 1977. /In Swedish/

Cliff, N. Adverbs as multipliers. In: Snider, J. & Osgood, C.E. Semantic differential technique. Chicago: Aldine, 1969. Pp 143-160.

Cooley, W.W. & Lohnes, P.R. Multivariate data analysis. New York: Wiley, 1971.

Dixon, W.J. (Ed.) Biomedical computer programs. Berkeley: University of California Press, 1975.

Guilford, J.P. Psychometric methods. New York: McGraw-Hill, 1954.

Lord, F.M. Some relations between Guttman's principal components of scale analysis and other psychometric theory. Psychometrika, 1958, 23 (4), 291-296.

Miller, G.A., Galanter, E. & Pribram, K.H. Plans and the structure of behavior. New York: Holt, 1960.

Nie, N.H., Hull, C.H., Jenkins, J.G., Steinbrenner, K. & Bent, D.H. Statistical package for the social sciences. (2nd ed.) New York: McGraw-Hill, 1975.

Quine, W.V. Epistemology naturalized. In: Royce, J.R. & Rozeboom, Wm.W. (Eds.) The psychology of knowing. New York: Gordon and Breach, 1972. Pp 9-23.

Ross, J.R. Adjectives as noun phrases. In. Reibel, D.A. & Schane, S.A. (Eds.) Modern studies in English. Englewood Cliffs, N.J.: Prentice-Hall, 1969. Pp 352-360.

Schank, R.C. Conceptual dependency: A theory of natural language understanding. Cognitive Psychology, 1972, 3 (4), 552-631.

Department of
Educational and
Psychological Research
School of Education
Malmö, Sweden

Abstract card

Bierschenk, B. A new approach to psychometric problems in the analysis of pre-numeric data. Didakometry (Malmö, Sweden: School of Education), No. 55, 1977.

This report suggests a new approach to quantification of the relations within a whole sentence and presents a system for the analysis of concepts by data processing (ANACONDA). The psychological model for ANACONDA is outlined. Panel assessments have been used for quantification. On the basis of the results obtained, relation patterns have been studied by means of a discriminant analysis. An attempt has been made to demonstrate the validity of the methodological approach used.

Keywords: Psycholinguistics, concept formation, cognition, interview data, psychometrics, information and documentation, regression analysis.

Reference card

Bierschenk, B. A new approach to psychometric problems in the analysis of pre-numeric data. Didakometry (Malmö, Sweden: School of Education), No 55, 1977.